

Holleran

Alt # 11
10/21/99 1642 Rush

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/978,217

DATE: 09/22/1999
TIME: 14:39:53

INPUT SET: S33430.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT: Benz, Christopher C.
6 Scott, Gary K.
7 Chang, Chuan-Hsiung
8
9 (ii) TITLE OF INVENTION: A New ETS-Related Gene Overexpressed in
10 Human Breast and Epithelial Cancers
11
12 (iii) NUMBER OF SEQUENCES: 38
13
14 (iv) CORRESPONDENCE ADDRESS:
15 (A) ADDRESSEE: Townsend and Townsend and Crew LLP
16 (B) STREET: Two Embarcadero Center, Eighth Floor
17 (C) CITY: San Francisco
18 (D) STATE: California
19 (E) COUNTRY: USA
20 (F) ZIP: 94111-3834
21
22 (v) COMPUTER READABLE FORM:
23 (A) MEDIUM TYPE: Floppy disk
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
27
28 (vi) CURRENT APPLICATION DATA:
29 (A) APPLICATION NUMBER: US 08/978,217
30 (B) FILING DATE: 25-NOV-1997
31 (C) CLASSIFICATION:
32
33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: US 60/031,504
35 (B) FILING DATE: 27-NOV-1996
36
37 (viii) ATTORNEY/AGENT INFORMATION:
38 (A) NAME: Hunter, Tom
39 (B) REGISTRATION NUMBER: 38,498
40 (C) REFERENCE/DOCKET NUMBER: 02307E-071110US
41
42 (ix) TELECOMMUNICATION INFORMATION:
43 (A) TELEPHONE: (415) 576-0200
44 (B) TELEFAX: (415) 576-0300
45
46

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/978,217DATE: 09/22/1999
TIME: 14:39:53

INPUT SET: S33430.raw

47 (2) INFORMATION FOR SEQ ID NO:1:
48

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 1116 base pairs

51 (B) TYPE: nucleic acid

52 (C) STRANDEDNESS: single

53 (D) TOPOLOGY: linear
5455 (ii) MOLECULE TYPE: cDNA
56
57

58 (ix) FEATURE:

59 (A) NAME/KEY: CDS

60 (B) LOCATION: 1..1116

61 (D) OTHER INFORMATION: /product= "human ESX"

62 /note= "epithelial-restricted with serine box (ESX)"
63
6465 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
66

67	ATG GCT GCA ACC TGT GAG ATT AGC AAC ATT TTT AGC AAC TAC TTC AGT	48
68	Met Ala Ala Thr Cys Glu Ile Ser Asn Ile Phe Ser Asn Tyr Phe Ser	
69	1 5 10 15	
70		
71	GCG ATG TAC AGC TCG GAG GAC TCC ACC CTG GCC TCT GTT CCC CCT GCT	96
72	Ala Met Tyr Ser Ser Glu Asp Ser Thr Leu Ala Ser Val Pro Pro Ala	
73	20 25 30	
74		
75	GCC ACC TTT GGG GCC GAT GAC TTG GTA CTG ACC CTG AGC AAC CCC CAG	144
76	Ala Thr Phe Gly Ala Asp Asp Leu Val Leu Thr Leu Ser Asn Pro Gln	
77	35 40 45	
78		
79	ATG TCA TTG GAG GGT ACA GAG AAG GCC AGC TGG TTG GGG GAA CAG CCC	192
80	Met Ser Leu Glu Gly Thr Glu Lys Ala Ser Trp Leu Gly Glu Gln Pro	
81	50 55 60	
82		
83	CAG TTC TGG TCG AAG ACG CAG GTT CTG GAC TGG ATC AGC TAC CAA GTG	240
84	Gln Phe Trp Ser Lys Thr Gln Val Leu Asp Trp Ile Ser Tyr Gln Val	
85	65 70 75 80	
86		
87	GAG AAG AAC AAG TAC GAC GCA AGC GCC ATT GAC TTC TCA CGA TGT GAC	288
88	Glu Lys Asn Lys Tyr Asp Ala Ser Ala Ile Asp Phe Ser Arg Cys Asp	
89	85 90 95	
90		
91	ATG GAT GGC GCC ACC CTC TGC AAT TGT GCC CTT GAG GAG CTG CGT CTG	336
92	Met Asp Gly Ala Thr Leu Cys Asn Cys Ala Leu Glu Glu Leu Arg Leu	
93	100 105 110	
94		
95	GTC TTT GGG CCT CTG GGG GAC CAA CTC CAT GCC CAG CTG CGA GAC CTC	384
96	Val Phe Gly Pro Leu Gly Asp Gln Leu His Ala Gln Leu Arg Asp Leu	
97	115 120 125	
98		
99	ACT TCC AGC TCT TCT GAT GAG CTC AGT TGG ATC ATT GAG CTG CTG GAG	432

RAW SEQUENCE LISTING PATENT APPLICATION US/08/978,217

DATE: 09/22/1999
TIME: 14:39:53

INPUT SET: S33430.raw

100	Thr Ser Ser Ser Ser Asp Glu Leu Ser Trp Ile Ile Glu Leu Leu Glu	
101	130 135 140	
102		
103	AAG GAT GGC ATG GCC TTC CAG GAG GCC CTA GAC CCA GGG CCC TTT GAC	480
104	Lys Asp Gly Met Ala Phe Gln Glu Ala Leu Asp Pro Gly Pro Phe Asp	
105	145 150 155 160	
106		
107	CAG GGC AGC CCC TTT GCC CAG GAG CTG CTG GAC GAC GGT CAG CAA GCC	528
108	Gln Gly Ser Pro Phe Ala Gln Glu Leu Leu Asp Asp Gly Gln Gln Ala	
109	165 170 175	
110		
111	AGC CCC TAC CAC CCC GGC AGC TGT GGC GCA GGA GCC CCC TCC CCT GGC	576
112	Ser Pro Tyr His Pro Gly Ser Cys Gly Ala Gly Ala Pro Ser Pro Gly	
113	180 185 190	
114		
115	AGC TCT GAC GTC TCC ACC GCA GGG ACT GGT GCT TCT CGG AGC TCC CAC	624
116	Ser Ser Asp Val Ser Thr Ala Gly Thr Gly Ala Ser Arg Ser Ser His	
117	195 200 205	
118		
119	TCC TCA GAC TCC GGT GGA AGT GAC GTG GAC CTG GAT CCC ACT GAT GGC	672
120	Ser Ser Asp Ser Gly Gly Ser Asp Val Asp Leu Asp Pro Thr Asp Gly	
121	210 215 220	
122		
123	AAG CTC TTC CCC AGC GAT GGT TTT CGT GAC TGC AAG AAG GGG GAT CCC	720
124	Lys Leu Phe Pro Ser Asp Gly Phe Arg Asp Cys Lys Lys Gly Asp Pro	
125	225 230 235 240	
126		
127	AAG CAC GGG AAG CGG AAA CGA GGC CGG CCC CGA AAG CTG AGC AAA GAG	768
128	Lys His Gly Lys Arg Lys Arg Gly Arg Pro Arg Lys Leu Ser Lys Glu	
129	245 250 255	
130		
131	TAC TGG GAC TGT CTC GAG GGC AAG AAG AGC AAG CAC GCG CCC AGA GGC	816
132	Tyr Trp Asp Cys Leu Glu Gly Lys Lys Ser Lys His Ala Pro Arg Gly	
133	260 265 270	
134		
135	ACC CAC CTG TGG GAG TTC ATC CGG GAC ATC CTC ATC CAC CCG GAG CTC	864
136	Thr His Leu Trp Glu Phe Ile Arg Asp Ile Leu Ile His Pro Glu Leu	
137	275 280 285	
138		
139	AAC GAG GGC CTC ATG AAG TGG GAG AAT CGG CAT GAA GGC GTC TTC AAG	912
140	Asn Glu Gly Leu Met Lys Trp Glu Asn Arg His Glu Gly Val Phe Lys	
141	290 295 300	
142		
143	TTC CTG CGC TCC GAG GCT GTG GCC CAA CTA TGG GGC CAA AAG AAA AAG	960
144	Phe Leu Arg Ser Glu Ala Val Ala Gln Leu Trp Gly Gln Lys Lys Lys	
145	305 310 315 320	
146		
147	AAC AGC AAC ATG ACC TAC GAG AAG CTG AGC CGG GCC ATG AGG TAC TAC	1008
148	Asn Ser Asn Met Thr Tyr Glu Lys Leu Ser Arg Ala Met Arg Tyr Tyr	
149	325 330 335	
150		
151	TAC AAA CGG GAG ATC CTG GAA CGG GTG GAT GGC CGG CGA CTC GTC TAC	1056
152	Tyr Lys Arg Glu Ile Leu Glu Arg Val Asp Gly Arg Arg Leu Val Tyr	

RAW SEQUENCE LISTING PATENT APPLICATION US/08/978,217

DATE: 09/22/1999
TIME: 14:39:54

INPUT SET: S33430.raw

```

153          340          345          350
154
155 AAG TTT GGC AAA AAC TCA AGC GGC TGG AAG GAG GAA GAG GTT CTC CAG      1104
156 Lys Phe Gly Lys Asn Ser Ser Gly Trp Lys Glu Glu Glu Val Leu Gln
157          355          360          365
158
159 AGT CGG AAC TGA      1116
160 Ser Arg Asn
161      370
162
163
164 (2) INFORMATION FOR SEQ ID NO:2:
165
166 (i) SEQUENCE CHARACTERISTICS:
167 (A) LENGTH: 371 amino acids
168 (B) TYPE: amino acid
169 (D) TOPOLOGY: linear
170
171 (ii) MOLECULE TYPE: protein
172
173 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
174
175 Met Ala Ala Thr Cys Glu Ile Ser Asn Ile Phe Ser Asn Tyr Phe Ser
176   1          5          10          15
177
178 Ala Met Tyr Ser Ser Glu Asp Ser Thr Leu Ala Ser Val Pro Pro Ala
179          20          25          30
180
181 Ala Thr Phe Gly Ala Asp Asp Leu Val Leu Thr Leu Ser Asn Pro Gln
182          35          40          45
183
184 Met Ser Leu Glu Gly Thr Glu Lys Ala Ser Trp Leu Gly Glu Gln Pro
185          50          55          60
186
187 Gln Phe Trp Ser Lys Thr Gln Val Leu Asp Trp Ile Ser Tyr Gln Val
188          65          70          75          80
189
190 Glu Lys Asn Lys Tyr Asp Ala Ser Ala Ile Asp Phe Ser Arg Cys Asp
191          85          90          95
192
193 Met Asp Gly Ala Thr Leu Cys Asn Cys Ala Leu Glu Glu Leu Arg Leu
194          100          105          110
195
196 Val Phe Gly Pro Leu Gly Asp Gln Leu His Ala Gln Leu Arg Asp Leu
197          115          120          125
198
199 Thr Ser Ser Ser Ser Asp Glu Leu Ser Trp Ile Ile Glu Leu Leu Glu
200          130          135          140
201
202 Lys Asp Gly Met Ala Phe Gln Glu Ala Leu Asp Pro Gly Pro Phe Asp
203          145          150          155          160
204
205 Gln Gly Ser Pro Phe Ala Gln Glu Leu Leu Asp Asp Gly Gln Gln Ala

```

RAW SEQUENCE LISTING PATENT APPLICATION US/08/978,217

DATE: 09/22/1999
TIME: 14:39:54

INPUT SET: S33430.raw

206		165		170		175
207						
208	Ser	Pro	Tyr	His	Pro	Gly
209				180		185
210						190
211	Ser	Ser	Asp	Val	Ser	Thr
212			195			200
213						205
214	Ser	Ser	Asp	Ser	Gly	Gly
215		210			215	
216						220
217	Lys	Leu	Phe	Pro	Ser	Asp
218	225				230	
219						235
220	Lys	His	Gly	Lys	Arg	Lys
221				245		250
222						255
223	Tyr	Trp	Asp	Cys	Leu	Glu
224			260			265
225						270
226	Thr	His	Leu	Trp	Glu	Phe
227		275			280	
228						285
229	Asn	Glu	Gly	Leu	Met	Lys
230		290			295	
231						300
232	Phe	Leu	Arg	Ser	Glu	Ala
233	305				310	
234						315
235	Asn	Ser	Asn	Met	Thr	Tyr
236			325			330
237						335
238	Tyr	Lys	Arg	Glu	Ile	Leu
239			340			345
240						350
241	Lys	Phe	Gly	Lys	Asn	Ser
242		355			360	
243						365
244	Ser	Arg	Asn			
245		370				
246						
247						
248	(2)	INFORMATION	FOR	SEQ	ID	NO:3:
249						
250	(i)	SEQUENCE	CHARACTERISTICS:			
25						

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/978,217

DATE: 09/22/1999
TIME: 14:39:55

INPUT SET: S33430.raw

Line

Error

Original Text